
ARTÍCULOS / ARTICLES

ALTERNATIVAS PARA LA MEDICIÓN DE LA DELINCUENCIA URBANA Y LA IDENTIFICACIÓN DE ZONAS CRIMINÓGENAS. NUEVOS INDICADORES BASADOS EN LA PRESENCIA DE POBLACIÓN

Alejandro Moreno Jiménez

Universidad de Guadalajara

alexmorejim@gmail.com

ORCID iD: <https://orcid.org/0000-0002-3988-1031>

Aurea Esther Grijalva Eternod

Universidad de Guadalajara

aurea.grijalva@academicos.udg.mx

ORCID iD: <https://orcid.org/0000-0001-8399-4247>

Recibido: 20/02/2021; Aceptado: 03/11/2022; Publicado: 21/12/2022

Cómo citar este artículo/citation: Moreno Jiménez, Alejandro y Grijalva Eternod, Aurea Esther (2022). Alternativas para la medición de la delincuencia urbana y la identificación de zonas criminógenas. Nuevos indicadores basados en la presencia de población. *Estudios Geográficos*, 83 (293), e121. <https://doi.org/10.3989/estgeogr.2022127.127>

Resumen: El objetivo de este trabajo es proponer formas alternativas de medición de las tasas de delincuencia en los entornos urbanos, comparando la utilización de la presencia de población y la medición tradicional, es decir, la población residencial. El análisis se realiza para tres tipos de delitos: violencia intrafamiliar, robo a persona y robo a negocio. Para ello, se sigue una metodología de análisis espacial de los datos de la Zona Metropolitana de Guadalajara (ZMG) en México. En los resultados se comprueba que, para algunos tipos de delitos, la presencia de población puede identificar con mayor precisión a la población en riesgo, lo que puede constituir una herramienta de política pública muy útil de cara a la prevención del delito, pues además de lograr una estimación más precisa del fenómeno, se pueden identificar las zonas generadoras y atractoras del crimen.

Palabras clave: seguridad ciudadana, criminología ambiental, zonas atractoras y generadoras de crimen, redes sociales, unidades económicas, presencia de población.

ALTERNATIVES FOR THE MEASUREMENT OF URBAN CRIME AND THE IDENTIFICATION OF CRIMINOGENIC ZONES. NEW INDICATORS BASED ON THE PRESENCE OF POPULATION

Abstract: The objective of this paper is to propose alternative ways of measuring crime rates in urban settings, comparing the use of the presence of population with the traditional measurement, that is, the residential population. The analysis is carried out for three types of crimes: domestic violence, personal robbery and business robbery. To achieve this, a methodology of spatial analysis of the data of the Guadalajara Metropolitan Area (ZMG), Mexico is conducted. The results show that, for some types of crime, the presence of population can more accurately identify the population at risk, which can constitute a very useful public policy tool for crime prevention, since in addition to achieving a more precise estimation of the phenomenon, the areas that generate and attract crime can be identified.

Key words: citizen security, environmental criminology, crime generators, crime attractors, social media, economic units, ambient population.

INTRODUCCIÓN

Una de las formas más extendidas de estimar el delito es a través de las tasas de incidencia o de prevalencia delictiva. Estas se suelen calcular como el número de víctimas o el número de delitos que ocurren en una zona determinada y en un tiempo concreto, dividido entre la cantidad de habitantes en ese mismo periodo y lugar. En esta operación, el denominador, es decir, la cantidad de habitantes, se suele obtener utilizando la población residencial observada en la unidad de análisis que se está utilizando. Sin embargo, esto puede suponer algunos problemas de precisión.

Es cierto que la población residencial puede ser el indicador más adecuado de la población en riesgo para ciertos tipos de delitos, por ejemplo, cuando se habla de robo a casa habitación y también cuando se están examinando unidades de análisis a un nivel macro como países o entidades federativas (Andresen, 2006). No obstante, es posible que, en áreas geográficas más pequeñas como los municipios o las colonias (barrios en España), la población residencial no sea la más adecuada para aproximarse a la población en riesgo, especialmente en el medio urbano que se caracteriza por una gran movilidad de las personas. Asimismo, en relación con ciertos delitos en los que el riesgo no se relaciona con el lugar de residencia, es muy posible que la población residencial no sea el mejor indicador (Andresen, 2011; Malleson & Andresen, 2014). En consecuencia, si los esfuerzos de prevención del crimen, especialmente en unidades de análisis pequeñas, se basan en tasas de delincuencia calculadas utilizando la población residencial, y estos no son adecuados, tales esfuerzos pueden no solo ser ineficaces para reducir el riesgo de eventos criminales, sino que además se puede estar propiciando un ineficiente uso de los recursos públicos.

Tomando en cuenta lo anterior, el objetivo del presente trabajo es proponer alternativas de medición que permitan informar de una manera más precisa, a los tomadores de decisiones en materia de seguridad ciudadana y prevención del delito, el lugar donde se concentran los delitos y en los que deberían centrar sus esfuerzos. El análisis se realiza en la Zona Metropolitana de Guadalajara (ZMG) en México, la cual, como se verá más adelante, es un caso de estudio de gran relevancia en función del problema de inseguridad creciente que ha experimentado en los últimos años. En concreto, se evalúa la utilización de datos provenientes de Twitter (tuits georreferenciados)¹ y del Directorio Estadístico

Anual de Unidades Económicas (DENUE) del Instituto Nacional de Estadística y Geografía (INEGI) como variables proxy a la presencia de población² y se realiza una comparación de tipo espacial con las estimaciones basadas en la población residencial. Con ello, se busca establecer si estas tres medidas de la población en riesgo permiten establecer patrones espaciales distintos y, en consecuencia, una mayor precisión en la estimación del delito. La delimitación temporal de este trabajo son cuatro meses: octubre, noviembre y diciembre de 2017, y enero de 2018.

Se considera que los tuits georreferenciados pueden ser una variable proxy³ de la presencia de población ya que representan el movimiento de las personas a través de las colonias de la ZMG. De igual forma, los datos de establecimientos y unidades económicas del DENUE nos servirán como un aproximado de la presencia de población ya que expresan la cantidad de empresas que hay en cada colonia y cuántas personas laboran en ellas. Para los datos del DENUE se parte del supuesto de que las personas asisten a los establecimientos a trabajar, pero también que los giros comerciales atraen personas que son susceptibles de ser víctimas de un crimen.

Así, la hipótesis de este trabajo es que, en algunos delitos, la población en riesgo se puede identificar con

México (66% de usuarios de redes sociales), detrás de Facebook (95%), WhatsApp (93%) y Youtube (72%). En este trabajo se utilizan datos de esta red social ya que los datos de las tres principales redes sociales del país no están disponibles libremente (Facebook), son confidenciales (Whatsapp) o su estructura no es compatible para un análisis espacial y estadístico (Youtube).

- 2 Para este trabajo, se utilizan los conceptos de presencia de población y población residencial. La presencia de población se entiende como la población móvil, es decir, la cantidad de personas que se puede esperar en un área en una determinada escala temporal, lo cual algunos también denominan población ambiente (Andresen, 2006), mientras que la población residencial es la población que se obtiene de censos y encuestas basadas en el hogar.
- 3 Es importante reconocer que el uso de Twitter ha generado un importante debate en las ciencias sociales y aún existen dudas relacionadas con su representatividad y con la existencia de sesgos sobre la población que representa, por lo que es importante tomar con cautela los resultados de investigaciones en los que se utilizan datos provenientes de esta red social (Blank, 2017). No obstante, la expansión de las redes sociales en los últimos tiempos hace que la utilización de estos datos sea cada vez más común dentro de la ciencia y que estas limitaciones vayan disminuyendo, tan es así que el uso de Big Data en el ámbito criminológico ha ido proliferando en tiempos recientes (Snaphaan & Hardyns, 2021).

1 La Asociación Mexicana del Internet (AMIPCI, 2016) enlista a Twitter como la cuarta red social más utilizada en

mayor precisión a través de la presencia de población en el ambiente representada por los tuits georreferenciados y la ubicación de establecimientos; mientras que, para otro tipo de delitos, puede ayudar más la población residencial reflejada en el Censo de Población y Vivienda del INEGI. El análisis se centrará en tres delitos en concreto: violencia intrafamiliar, robo a persona y robo a negocio. En específico, se espera que la población residencial tenga una mayor correlación con la violencia intrafamiliar (H_1), ya que este delito suele acontecer en el núcleo familiar; también se espera que la presencia de población representada por tuits tenga una mayor correlación con robos a persona (H_2) pues es posible que indiquen con mayor precisión el movimiento de las personas entre las colonias de la ZMG; y, por último, se espera que la presencia de población representada por establecimientos obtenga una mayor correlación con el delito de robo a negocio (H_3).

Es importante destacar que, además de la posible precisión que se pretende alcanzar, existen algunas ventajas de utilizar la presencia de población a través de tuits y establecimientos comerciales sobre la utilización de la población residencial. Algunas de estas ventajas son: 1) Twitter es una fuente de información gratuita, 2) las Encuestas Intercensales no son estadísticamente significativas a un nivel tan detallado como lo es una zona metropolitana y las manzanas que la conforman, no obstante, la geolocalización ayuda a que la información de redes sociales sea granular espacial y temporalmente, 3) la periodicidad, ya que el INEGI elabora un censo poblacional cada 10 años, esto tiene como consecuencia que muchos datos se encuentren desactualizados y sean imprecisos; en cambio, los datos que arroja la API de twitter especifican la hora y día que se publicó dicho tuit y la información del DENU se actualiza cada cuatro meses; 4) los datos tradicionales, como los censos u otros estudios basados en el hogar, son temporalmente estáticos y no son capaces de capturar información adecuada sobre las actividades y el comportamiento fuera del hogar.

Aunado a lo anterior, el uso de datos provenientes de redes sociales en las decisiones públicas ha ido incrementándose, particularmente en temas de movilidad. Un ejemplo de ellos es el esfuerzo realizado por el Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación (INFOTEC) en coordinación con el Centro de Investigación en Ciencias de Información Geoespacial (CentroGEO) y el Consejo Nacional de Ciencia y Tecnología (2020), en el que se mide la movilidad en las entidades federativas de México a través de tres redes sociales: Facebook,

Google y Twitter y que ha abonado a la estrategia gubernamental de disminución de propagación de la enfermedad COVID-19 en México. Asimismo, a nivel global, los tuits geolocalizados han mostrado ser una buena aproximación para medir los patrones de movilidad (Hawelka et al. 2014).

ANTECEDENTES TEÓRICOS

Espacio, ambiente y criminalidad

Existen cuatro dimensiones básicas para entender el crimen: la legal, la de la víctima, la del delincuente y la del lugar o situación (Brantingham & Brantingham, 1981). La Criminología Ambiental se concentra en la cuarta dimensión y consiste en una familia de teorías criminológicas que comparten un común interés en eventos criminales y las circunstancias inmediatas en las cuales ocurren (Wortley & Mazerolle, 2008) preguntándose fundamentalmente dónde y cuándo ha ocurrido el crimen (Brantingham & Brantingham, 1981).

De acuerdo con Wortley y Mazerolle (2008), las principales premisas de la Criminología Ambiental son: 1) el comportamiento criminal se encuentra significativamente influenciado por el entorno en el que se encuentra; 2) los delitos tienden a concentrarse en ciertos lugares y momentos concretos cometidos por personas que se dirigen o se encuentran en un determinado lugar; y, 3) la comprensión de las características sociales y espaciales de los eventos criminales son elementos de utilidad para la investigación, control y prevención de los mismos. Esto es importante ya que el análisis del delito por parte de las autoridades competentes debería realizarse a un nivel espacial y temporal tratando de identificar patrones y tendencias del crimen (Wortley & Mazerolle, 2008).

Las tasas de crimen se ven afectadas por una convergencia en tiempo y espacio de ofensores motivados, objetivos adecuados y guardianes competentes contra el crimen (Cohen & Felson, 1979). Con base en lo anterior, dentro de la Criminología Ambiental se han desarrollado los conceptos de áreas generadoras de crimen, áreas atractoras de crimen y áreas neutras. Las zonas generadoras de crimen son áreas que atraen grandes cantidades de personas por motivos que no están relacionados con alguna motivación criminal, por ejemplo, centros comerciales o restaurantes. Por su parte, las zonas atractoras son áreas con oportunidades conocidas de crimen donde los criminales son atraídos por las oportunidades conocidas de particulares tipos de crimen, por ejemplo: áreas de

prostitución, mercados de drogas, etc. Finalmente, las áreas neutras no atraen a los delincuentes con intención porque esperan cometer un delito, ni producen grandes oportunidades delictivas, pero sí experimentan crímenes ocasionales por delincuentes sobre todo locales (Brantingham & Brantingham, 1995).

Las áreas atractoras, generadoras y neutras se pueden identificar a través de *hotspots* que se ven influidos por dónde viven las personas, su desplazamiento en la ciudad y cómo pasan su tiempo. El desplazamiento del delito depende del tipo de área que representa cada *hotspot*, así, es improbable que la intervención en áreas identificadas como generadoras de crimen provoque el desplazamiento delictivo ya que los delitos que allí ocurren son oportunistas; sin embargo, los desplazamientos en áreas atractoras de crimen son mucho más probables por tres razones. En primer lugar, es probable que la actividad delictiva se desplace hacia el vecindario que rodea a la zona atractora si hay objetivos llamativos o víctimas cercanas. En segundo lugar, es probable que la actividad criminal que no puede desplazarse al vecindario que rodea a la zona atractora original se desplace a otros nodos de atracción importantes. Tercero, la actividad criminal que no puede desplazarse al vecindario que rodea a la zona atractora y no se desplace a otro atractor importante del delito, puede ser desplazada al vecindario de origen del delincuente (Brantingham & Brantingham, 2008).

Criminología y *Big Data*

La gestión y el análisis de grandes volúmenes de datos (*Big Data*) han sido de gran utilidad para contribuir a la solución de algunos problemas sociales. Específicamente, en temas de prevención del crimen urbano, el *Big Data* también ha tenido una gran influencia en años recientes, especialmente porque se ha considerado que para que exista inteligencia policial se necesita información que ayude a predecir, a vigilar el internet y que pueda tener retroalimentación en tiempo real (Cheng y Chen, 2021). Al respecto, Chan y Bennett (2015) sostienen que hay dos áreas principales en el que este tipo de datos se han utilizado dentro de la investigación del fenómeno criminal: como fuente de datos y como una herramienta predictiva para guiar estrategias de prevención o para orientar las decisiones en el área de justicia penal.

Para ejemplificar lo anterior basta mencionar algunos casos. En primer lugar, la policía de la ciudad de Chicago ha aplicado técnicas como el *Machine Learning*⁴ y el análisis predictivo de datos para prevención del crimen a través de una herramienta de predicción geográfica que utiliza el modelado de información para predecir el riesgo en zonas específicas de la ciudad. De igual forma, existe una herramienta realizada por IBM llamada *Coplink*, en la cual se agrupan datos sobre arrestos, fotografías policiales, datos de ubicación e información sobre pandillas en un tablero donde la policía puede ver y compartir información. Además, IBM ha trabajado de la mano con la policía de Manchester con un software predictivo que ha ayudado a reducir los robos de vehículos a través de datos históricos. Finalmente, *PredPol* es un programa que sirve para determinar aquellas zonas de alguna ciudad en las que existe una mayor probabilidad de que exista un crimen y es utilizada por cuerpos policiales en distintas ciudades de Estados Unidos y Europa.

También existen algunos trabajos científicos que han hecho uso del *Big Data*, para aportar evidencia empírica a las teorías criminológicas. En este sentido, Williams et al (2016) han relacionado la teoría de las ventanas rotas con la información que deriva de los tuits georreferenciados para estimar patrones de la criminalidad, encontrando que las menciones de Twitter de indicadores relacionados en su contenido con deterioro ambiental, están significativamente asociadas con varios tipos de delitos. Otro trabajo relacionado con este aspecto es el de Wang et al. (2012) quienes utilizan datos de Twitter para predecir incidentes criminales. El enfoque de este trabajo se basa en el análisis semántico automático y la comprensión del lenguaje natural de tuits, combinado con la reducción de la dimensionalidad a través de la asignación de Dirichlet y la predicción a través de modelos lineales.

De igual forma, Traunmuller et al. (2014) proponen un método para comprobar cuantitativamente teorías del crimen urbano mediante datos de telecomunicaciones y de delitos en Londres. Utilizan conteos de peatones registrados por datos de telecomunicaciones, para determinar si al existir correlación con los registros de datos abiertos de delincuencia, se puede revelar en dónde es aplicable estas teorías. Particularmente, encuentran que algunas variables como la diversidad de la población y la proporción de hombres jóvenes son relevantes espacialmente. Wang et al. (2016) es otro ejemplo del uso de *Big Data* para inferencia de tasas de delincuencia con datos de flujos de taxis. En este trabajo se encontró que utilizando este

4 El *Machine Learning* es un área de la inteligencia artificial que se basa en algoritmos que pueden aprender de los datos sin depender de la programación basada en reglas.

nuevo tipo de información se mejora la inferencia de la tasa de delincuencia en la ciudad de Chicago.

Es claro entonces que el *Big Data* ha aportado y seguirá aportando información útil para la prevención del crimen ya que es un mercado en expansión en el que cada vez se interesan más agentes tanto del sector público como del sector privado. Y, particularmente, la información que deriva de redes sociales es una estructura que parece que cada día tiene mayor alcance en la seguridad ciudadana, tal y como se ha mostrado en diversos trabajos.

No obstante, hay que resaltar que el uso de Big Data también implica algunos retos y limitaciones para el conocimiento criminológico que deben tenerse en cuenta y que obligan a los investigadores a ser más rigurosos desde el punto de vista metodológico (Pranik et al., 2017). Entre las limitaciones más importantes se encuentran que los datos no fueron creados para fines de investigación y la existencia de sesgos de muestreo y de selección que derivan del uso de algunas redes sociales (Snaphaan & Hardyns, 2021). Sin embargo, como proponen Chan y Bennett (2015), es solo hasta que existan más trabajos que utilicen de manera efectiva el *Big Data* para contrastar teorías y generar conocimiento criminológico, cuando se sabrá si se transformarán las prácticas tradicionales.

Presencia de población y estimación del delito

La idea de que la estimación de los delitos utilizando la población residencial puede tener muchos sesgos para comprender la población que se encuentra en riesgo no es nueva, e incluso, existen ya diversos trabajos que han abordado esta situación de manera empírica.

Malleson y Andresen (2014), por ejemplo, se plantearon como objetivo destacar las áreas que sufren altos índices de delincuencia utilizando tanto población residencial como presencia de población en la estimación de la población en riesgo. Para ello, utilizaron tuits como variable proxy para medir la presencia, partiendo de que la población residencial es inadecuada como medida de población en riesgo de delito, ya que muchos delitos involucran personas en movimiento como asaltos, robo a automóviles, etc. Si bien, los mismos autores reconocen las limitaciones del uso de tuits, su trabajo demuestra que diferentes patrones espaciales de los índices de delincuencia emergen cuando se usan la población residencial y la presencia de población.

Por otra parte, en Trinidad et al. (2019) se analiza la concentración espacial de la delincuencia juvenil en

el área urbana de la ciudad de Bilbao en la Comunidad Autónoma del País Vasco, España. Sus resultados muestran que los delitos se encuentran concentrados en espacios en los que se ubican establecimientos de giros muy específicos, destacando el comercio, los bancos y la comida rápida. Además, se infiere que las áreas en el centro de la ciudad serían espacios generadores de delito, ya que atraen a personas cuya intención primaria no es delinquir.

Asimismo, Andresen (2006) investiga la actividad criminal desde el punto de vista espacial en Vancouver, empleando la teoría de la desorganización social, la teoría de las actividades rutinarias y las medidas de delincuencia. Para esto, la presencia de población es calculada como un promedio de 24 horas de una población en una unidad espacial para capturar la población en riesgo, y encuentran que existe un fuerte respaldo para la teoría de las actividades rutinarias en el espacio y el uso de las poblaciones ambientales al calcular las tasas de criminalidad. También, en Andresen y Jenion (2008) se analiza el uso de la presencia de población para informar mejor las iniciativas de prevención del delito. Los resultados indican que la presencia de población tiene utilidad para los tres niveles de prevención del delito (primario, secundario y terciario).

Adicionalmente, Andresen (2011) utiliza la presencia de población para el cálculo de las tasas de criminalidad, como una medida alternativa a la de la población en riesgo, al considerar que al calcular tasas de delincuencia con la población residencial surge un problema ya que las personas dejan sus límites censales (hogares) durante el día, esto trae como consecuencia que las tasas de delincuencia basados en población residencial podrían no ser precisas. Para este trabajo se utilizó la presencia de población como una estimación promedio de veinticuatro horas de la población en una unidad espacial disponible a través del Laboratorio Nacional *Oak Ridge*. En el mismo sentido, en Ristea et al. (2017) se encuentra correlación espacial entre algunos tipos de conductas antisociales (como vandalismo, tomar alcohol en vía pública, tirar basura o pequeños robos) y tuits con contenido violento en las proximidades de dos estadios en Manchester, Inglaterra. Esta relación se explica, con base en Andresen y Linning (2012), por el uso de suelo y la movilidad de las personas, así como la existencia de pubs, paradas de transporte y centros de ocio.

Todo lo anterior muestra que existe evidencia previa sobre la correlación espacial entre la presencia de población, medida a través de los tuits, y la distribución de algunos tipos de delito en concreto, así como

el potencial de usar tuits geolocalizados para identificar patrones espaciales de los distintos delitos, lo que motiva al presente estudio en donde se pretende generar evidencia en un contexto donde esta relación no ha sido prácticamente explorada.

METODOLOGÍA

Variables

Para el presente trabajo se utilizaron las siguientes fuentes de datos:

- Las denuncias a nivel colonia para distintos tipos de delito generados por la FGEJ.
- El grado de marginación por colonia calculado por el Consejo Nacional de Población (CONAPO).
- La población residencial se calculó con la cantidad de personas que habitan por colonia de acuerdo con el Censo de Población y Vivienda 2010 del INEGI.
- Para la presencia de población se utilizaron dos variables: los tuits georreferenciados por colonia⁵ y los datos sobre establecimientos y usos de tierra⁶ del DENUE del INEGI. Los establecimientos se clasificaron, de acuerdo con Kinney et al. (2008), en: centros comerciales; bares, antros y centros nocturnos; restaurantes; estacionamientos; centros de transporte; escuelas, universidades y colegios; y edificios de centros culturales y recreacionales.

Por otro lado, para extraer los tuits georreferenciados se utilizó el Software R Project y se trabajó con la paquetería TwitterR, la cual provee una interfaz a la API web de Twitter. Para el código se define un centroide de la ZMG y un buffer que abarque el total de colonias de la ZMG. El centroide se definió en el punto con coordenadas -103.356358,20.664462, dicho centroide se calculó a través del software Qgis. El Buffer generado tiene 30 km de radio y abarca el total de las colonias de la ZMG.

Del total de tuits, se filtraron aquellos que cuentan con coordenadas, es decir, tuits georreferenciados, y a través del software Qgis se ubicó la cantidad de tuits

que se encuentran dentro del polígono que representa cada colonia, del total de tuits extraídos durante los cuatro meses estudiados, solo el 2,5% cuenta con coordenadas. Del total de tuits que cuentan con coordenadas, el 52% proviene de publicaciones en Instagram vinculadas a Twitter, el 35% de publicaciones en Foursquare vinculadas a Twitter y el 1% provienen de publicaciones en SwarmApp vinculadas a Twitter, esto da un total de 88,3% de tuits georreferenciados que provienen de publicaciones en otras redes sociales. Por otro lado, el 4,1% son tuits georreferenciados desde aplicaciones de Twitter para celulares con sistema operativo Android, IOs y Windows, es decir, una publicación que un usuario realizó directamente y en el que precisó su ubicación, sin provenir de otras redes sociales o aplicaciones, lo que se podría definir como tuits puros (no bots)

Es importante señalar que se detectaron colonias que contaban con el valor cero en las distintas variables consideradas, especialmente en los datos de denuncias de delitos, tuits y grado de marginación. De las 1348 colonias de la ZMG existieron 371 en las que no se denunciaron crímenes en el lapso de estudio, esto representa el 27,5% del total de colonias, existen también 282 colonias donde se reportaron delitos, pero no se presentaron tuits georreferenciados (21% del total de colonias), también es importante considerar que 231 colonias no presentaron información de grado de marginación⁷.

Por lo anterior, para el cálculo de las tasas de delincuencia considerados en este trabajo y para las variables utilizadas como denominadores (censo, tuits y DENUE) que tenían valores de cero, los mismos fueron sustituidos por el valor uno, de este modo se pueden obtener valores reales aproximados para calcular dichas tasas. Este procedimiento de imputación puede tener como consecuencia una sobreestimación de las tasas de criminalidad ya que afecta a la varianza y a la media. Sin embargo, de no realizarse este procedimiento la presencia de población se podría subestimar (Maltz, 1999). El método de imputación usado es el de reemplazo univariado, en el que se sustituyen

5 Extraídos a través de la API de twitter

6 Los datos sobre establecimientos y usos de tierra del DENUE es una base de datos sobre empresas y organizaciones establecidas en México, para este estudio se consideran aquellos establecimientos cuyas coordenadas caigan dentro del área de las colonias de la ZMG

7 Si bien lo anterior puede constituir una limitante de los datos, es un problema que se ha observado en otros trabajos realizados en países en los que el uso de redes sociales es mayor; e incluso, este problema se esperaría aún si se consideraran periodos de tiempo más prolongados (Malleon y Andresen, 2016). De hecho, el problema no solo tiene que ver con los tuits, pues en el caso de los delitos, Andresen (2011) observó que el robo tenía valores de cero en el 37% de las unidades consideradas para un periodo de tres años.

los ceros por un valor constante entre 0 y el valor límite detectado (Lubbe et al. 2021). Debido a que las variables son discretas se justifica la sustitución de ceros por unos. El porcentaje sustituido en las variables es el siguiente: presencia de población con establecimientos 11,6%, presencia de población con tuits 39,4%, población residencial 10,2% y denuncias de delitos 27,3%. Aunque el porcentaje de sustitución es alto para las variables de presencia de población con tuits y denuncias de delitos, el resultado no presenta un cambio alto en la media y la varianza de dichas variables. En el caso de presencia de población con tuits la varianza disminuye -0,06% y la media crece un 0,8%. En el caso de las denuncias de delitos la varianza disminuye 1% mientras que la media crece 2%.

Estrategia de análisis

1. En el análisis se realizaron los siguientes pasos:
2. Estadística descriptiva para facilitar el entendimiento del contexto en el que se elabora este trabajo
3. Cálculo de correlaciones entre el total de delitos y las distintas variables que describen la población residencial y la presencia de población. También se realizó el cálculo de correlaciones con el grado de marginación con el fin de conocer si en realidad la brecha digital se ve reflejada en las variables utilizadas. Asimismo, se calcularon correlaciones entre los distintos tipos de delito y las distintas medidas de población para poder determinar la forma más adecuada de estimar las tasas para el total y por tipo de delito
4. Análisis espacial comenzando por un análisis de la distribución geográfica de las variables consideradas. Esto con el objetivo de realizar un análisis exploratorio preliminar sobre las distribuciones de las distintas variables. A partir de entonces se calculan las tasas de delincuencia utilizando los distintos tipos de población (residencial o presencia) como denominador

Para la comparación de las tres medias, las tasas de delincuencia se calcularon como:

$$TDR_i = \frac{dendel_i}{pobres_i}$$

Donde:

TDR_i = Tasa de delincuencia basada en población residencial en colonia i

$dendel_i$ = Denuncias de delitos en colonia i

$pobres_i$ = Población residencial en colonia i

$$TDAD_i = \frac{dendel_i}{pobad_i}$$

Donde:

$TDAD_i$ = Tasa de delincuencia basada en presencia de población representada por establecimientos y unidades económicas en colonia i

$dendel_i$ = Denuncias de delitos en colonia i

$pobad_i$ = Cantidad de establecimientos y unidades económicas en colonia i

$$TDAT_i = \frac{dendel_i}{pobat_i}$$

Donde:

$TDAT_i$ = Tasa de delincuencia basada en presencia de población representada por tuits georreferenciados en colonia i

$dendel_i$ = Denuncias de delitos en colonia i

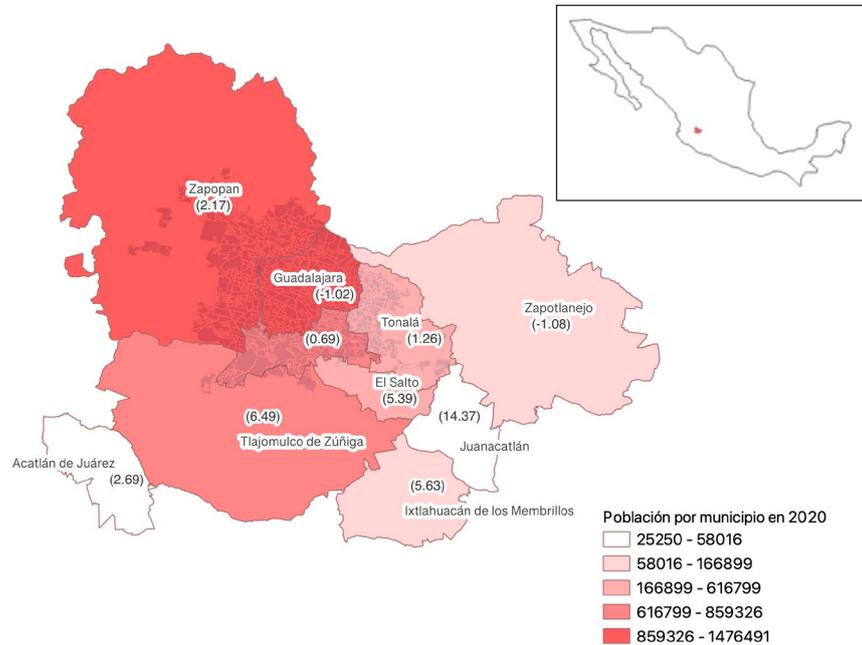
$pobat_i$ = Cantidad de tuits georreferenciados en colonia i

De igual forma, se trabajó con el estadístico Getis-Ord G_i^* (Getis & Ord, 1992) para generar *hotspots* (*clusters*). El estadístico G_i^* opera al examinar cada ubicación i (en este caso la ubicación es cada colonia) junto con sus ubicaciones vecinas j. Se calcula si el número total (o tasa) de ocurrencias en i y j es mayor o menor de lo que se esperaría por casualidad en comparación con las ubicaciones circundantes hasta una distancia d alejada de i. Si se encuentra una diferencia, se supone que las áreas i y j están asociadas y son diferentes a su entorno, es decir, es un *hotspot* o un *coldspot*. Un beneficio de este procedimiento sobre otros es su capacidad de proporcionar un puntaje Z y un valor p para cada ubicación. Estos indican si la hipótesis nula (no asociación entre el nivel de criminalidad en las áreas i y j) puede ser rechazada. Por lo tanto, es posible identificar *hotspots* estadísticamente significativos (Malleon & Andresen, 2016).

Caso de Estudio

El área de estudio es la ZMG la cual se encuentra en el poniente de México, en la entidad federativa de Jalisco. De acuerdo con el INEGI, durante 2020 su población fue de 5,27 millones de habitantes. Esto la convierte en la tercera Zona Metropolitana con mayor

FIGURA 1.
ZONA METROPOLITANA DE GUADALAJARA



Nota: En paréntesis se muestra el crecimiento poblacional anual de 2015 a 2020

población en México. Los diez municipios que conforman la ZMG son: Acatlán de Juárez, Guadalajara, Ixtlahuacán de los Membrillos, Juanacatlán, El Salto, Tlajomulco de Zúñiga, San Pedro Tlaquepaque, Tonalá, Zapopan y Zapotlanejo⁸. En la figura 1 se aprecia la población en los municipios que la integran, así como su crecimiento poblacional anual de 2015 a 2020. Por otro lado, como se señalaba en la introducción, la ZMG es un área de estudio particularmente interesante debido a su dinámica delictiva y el incremento de inseguridad ciudadana que se ha presentado en los últimos años. De acuerdo con datos del Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública, durante 2019 los delitos de robo en Jalisco incrementaron 177% respecto a 2015, destacando el incremento de robo a persona (575%), robo a negocios (435%) y robo a casa habitación (43%). Específicamente, en la ZMG: 1) en el municipio de Zapopan, el robo a persona creció 1089% en 2019 respecto a 2015 y en el municipio de Tonalá 1259%; 2) en el municipio de San

Pedro Tlaquepaque el robo a negocio creció 1082% en el mismo periodo; y 3) en el municipio de Guadalajara y en el de Zapopan creció el robo a casa habitación 179% y 232% respectivamente. También es importante considerar que durante 2019 el 40% de los robos del total de Jalisco se presentaron en Guadalajara, el 19% en Zapopan, el 10,5% en Tlaquepaque y el 8,7% en Tonalá, en total, los robos presentados en estos 4 municipios representaron el 69,4% del total de robos en Jalisco durante 2019, mientras que en 2015 esta proporción fue de 43,2%.

En la figura 2 se puede apreciar que en diciembre se reportaron la menor cantidad de delitos y que en octubre se registró la mayor. Asimismo, Guadalajara fue el que mayor cantidad de delitos registró y Tonalá el que menos. En la figura 3 se puede apreciar que el robo a vehículos fue el tipo de delito más reportado, seguido de amenazas y violencia intrafamiliar, mientras que robo a casa habitación fue el delito menos reportado.

Por otro lado, dado que se asume que los datos extraídos de Twitter representan la movilidad de las personas se realizó un ejercicio para dar robustez a

8 En este trabajo sólo se consideraron los municipios de Guadalajara, Zapopan, Tonalá y Tlaquepaque ya que son los únicos en los que la Fiscalía General del Estado de Jalisco (FGEJ, 2021) reporta datos de delitos por colonia.

FIGURA 2.
CANTIDAD DE DENUNCIAS REGISTRADAS POR TIPO DE DELITO EN LA ZMG DE OCTUBRE DE 2017 A ENERO DE 2018

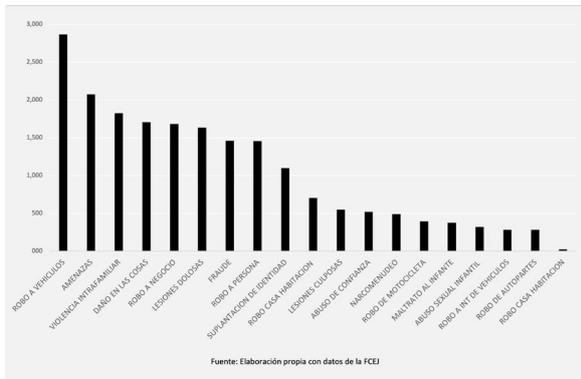
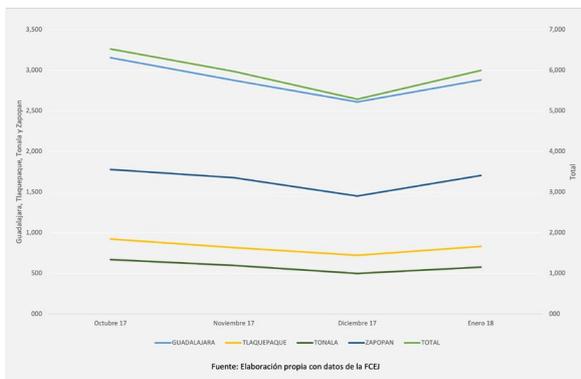


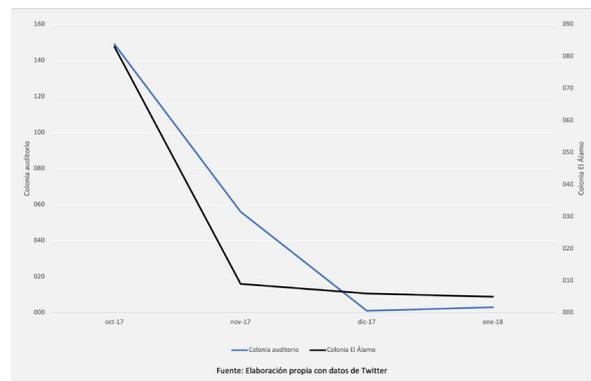
FIGURA 3.
DENUNCIAS POR MES Y MUNICIPIO EN LA ZMG



esta presunción basado en Osorio (2020). Los resultados ayudaron a comprobar que los tuits geolocalizados son una fuente eficaz para analizar los patrones espaciales en eventos sociales. Se eligieron la colonia Auditorio y la colonia El Álamo, las cuales son sede de dos eventos importantes en la ZMG durante el mes de octubre que son las Fiestas de Octubre y la Expo Ganadera, respectivamente. Estos eventos son muy

concurridos por los habitantes de la ZMG y dichas colonias tienen la característica de que, fuera del mes de octubre, no tienen otros atractores que provoquen movilidad de personas hacia dichas colonias. En la figura 4 se puede apreciar que efectivamente durante el mes de octubre de 2017 se presentaron en ambas colonias una mayor cantidad de tuits a diferencia de los meses de noviembre y diciembre de 2017 y enero de 2018. Este ejercicio confirma que, a pesar de las limitaciones que tiene el uso de tuits y que se han reconocido expresamente tanto en este como en otros trabajos, pueden ser un *proxy* adecuado para establecer la presencia de la población en un espacio.

FIGURA 4.
TUIITS IDENTIFICADOS POR COLONIA



RESULTADOS

Análisis correlacional

En la tabla 1 se observa la matriz de correlaciones entre las variables analizadas. Si bien se encontró que existe una correlación negativa y estadísticamente significativa entre el grado de marginación, las denuncias y los tuits generados, la misma es muy baja (-0,11 y -0,096 respectivamente), por lo que es posible asumir que ello no constituye una limitación importante

TABLA 1
MATRIZ DE CORRELACIONES

	Denuncias	Tuits	Población residencial	Establecimientos	Marginación
Denuncias	1				
Tuits	0,5611***	1			
Población residencial	0,4325***	0,088**	1		
Establecimientos	0,836***	0,5465***	0,3875***	1	
Marginación	-0,1127***	-0,0964**	-0,0283	-0,0426	1

Fuente: Elaboración propia con datos extraídos de la API de twitter, Censo de Población 2010, DENUY y FGEJ. * p<.05; ** p<.01; *** p<.001

TABLA 2
CORRELACIONES ENTRE TIPOS DE POBLACIÓN Y TIPOS DE DELITO

Tipo de delito	Tuits	Establecimientos	Población residencial
Robo A Negocio	0,5897***	0,7867***	0,3017***
Robo A Persona	0,6295***	0,7775***	0,2758***
Violencia Intrafamiliar	0,1634***	0,5342***	0,4571***

Fuente: Elaboración propia con datos extraídos de la API de twitter, Censo de Población 2010, DENUe y FGEJ. * p<,05; ** p<,01; *** p<,001

TABLA 3
CORRELACIONES ENTRE TASAS DE DELINCUENCIA

	Tasa con tuits	Tasa con población residencial	Tasa con establecimientos
Tasa con tuits	1		
Tasa con población residencial	0,1649***	1	
Tasa con establecimientos	0,4668***	0,0956***	1

Fuente: Elaboración propia con datos extraídos de la API de twitter, Censo de Población 2010, DENUe y FGEJ. * p<,05; ** p<,01; *** p<,001

para el presente análisis. Se puede apreciar también que la correlación entre la población residencial y los tuits también es baja (0,088), por lo que, al parecer, los tuits otorgan información adicional que explica el delito. Mientras que la correlación entre tuits y establecimientos es mayor (0,547), lo cual es esperable y deseable pues ambas medidas reflejan la presencia de población.

Como se puede apreciar en la tabla 2 y coincidiendo con lo planteado en las hipótesis; la población residencial obtuvo una mayor correlación con la violencia intrafamiliar, la presencia de población representada por tuits obtuvo mayor correlación con robos a persona y la presencia de población representada por establecimientos y empresas generó una correlación mayor con robos a negocios.

Por otro lado, en la tabla 3 se puede apreciar que existe una mayor correlación entre las tasas de delincuencia calculadas con los dos distintos tipos de presencia de población que con la tasa de delincuencia calculada con población residencial. Ambas correlaciones son estadísticamente significativas. Este era el resultado esperado ya que ambas tasas representan con distintos datos el movimiento de las personas entre colonias y ello también nos confirma que los tuits pueden utilizarse para aproximarnos a la presencia de población, a pesar de las limitaciones antes mencionadas.

Análisis espacial

Análisis Exploratorio de Datos Espaciales

En la figura 5 se aprecian los tuits por colonia recolectados durante los meses de estudio, se pueden apreciar colonias con mayor intensidad de tuits publicados, específicamente en la zona centro y occidente. En la figura 6, que muestra el total de los delitos considerados en este trabajo, también se puede apreciar que la mayor concentración se encuentra en la zona centro. Esto implica una relación entre ambas variables.

En la figura 5 se muestra también la distribución por colonia de establecimientos por cantidad de personas que laboran en ellos, apreciando una concentración en la zona centro, pero también en algunas colonias al occidente de la ciudad, al igual que ocurre con los delitos. También se aprecia la distribución de la población residencial por colonia, observando una distribución homogénea, lo cual es esperable dado que la residencia de las personas no tiene que ver con su movilidad, pero ello también muestra visualmente que no hay un comportamiento parecido entre población residencial y la ocurrencia de delitos, a diferencia de lo que ocurre con las dos medidas de presencia de población.

La figura 7 presenta la tasa de delincuencia por colonia calculada utilizando como denominador la pre-

FIGURA 5.
POBLACIÓN RESIDENCIAL, PRESENCIA DE POBLACIÓN (TUITS), PRESENCIA DE POBLACIÓN (ESTABLECIMIENTOS)

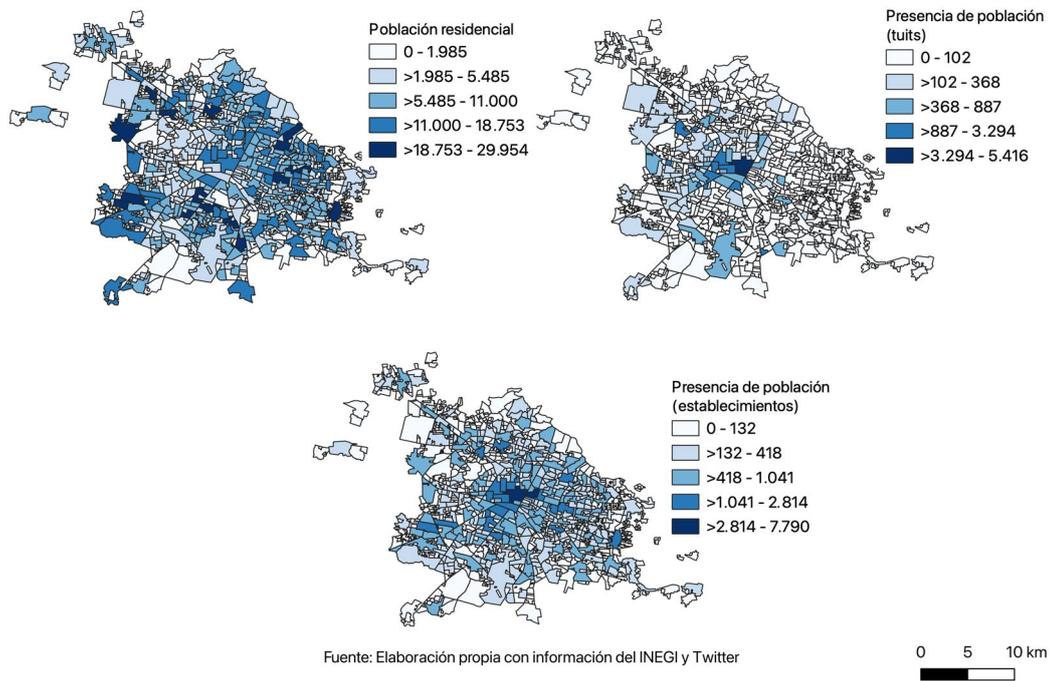


FIGURA 6.
TOTAL DE DELITOS, ROBOS A PERSONA, ROBOS EN ESTABLECIMIENTOS Y VIOLENCIA INTRAFAMILIAR

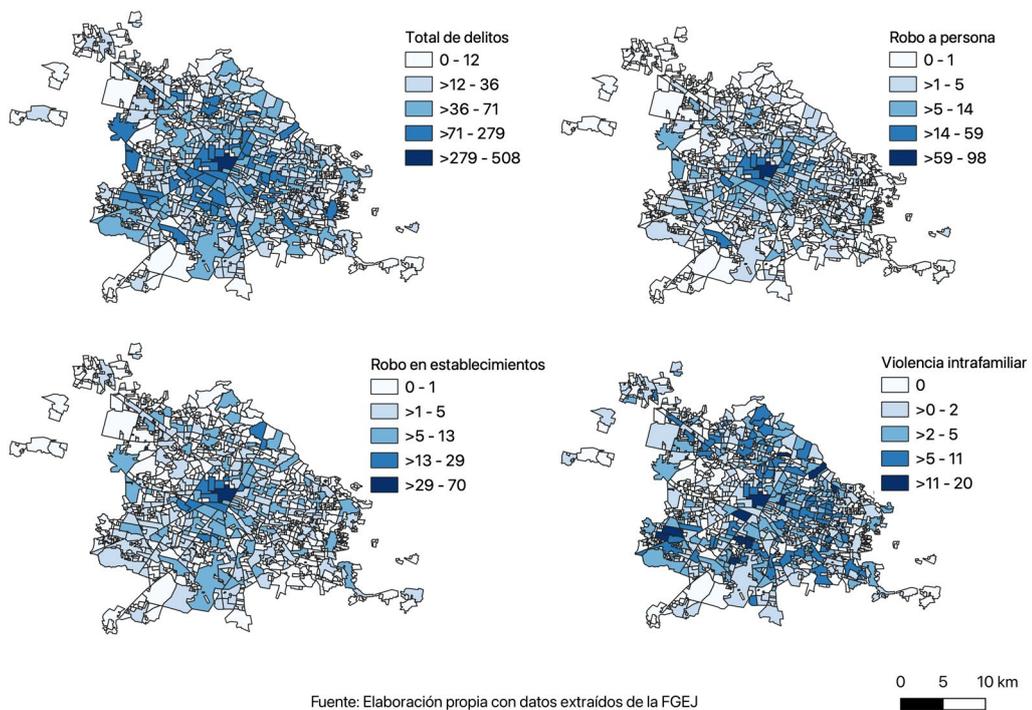
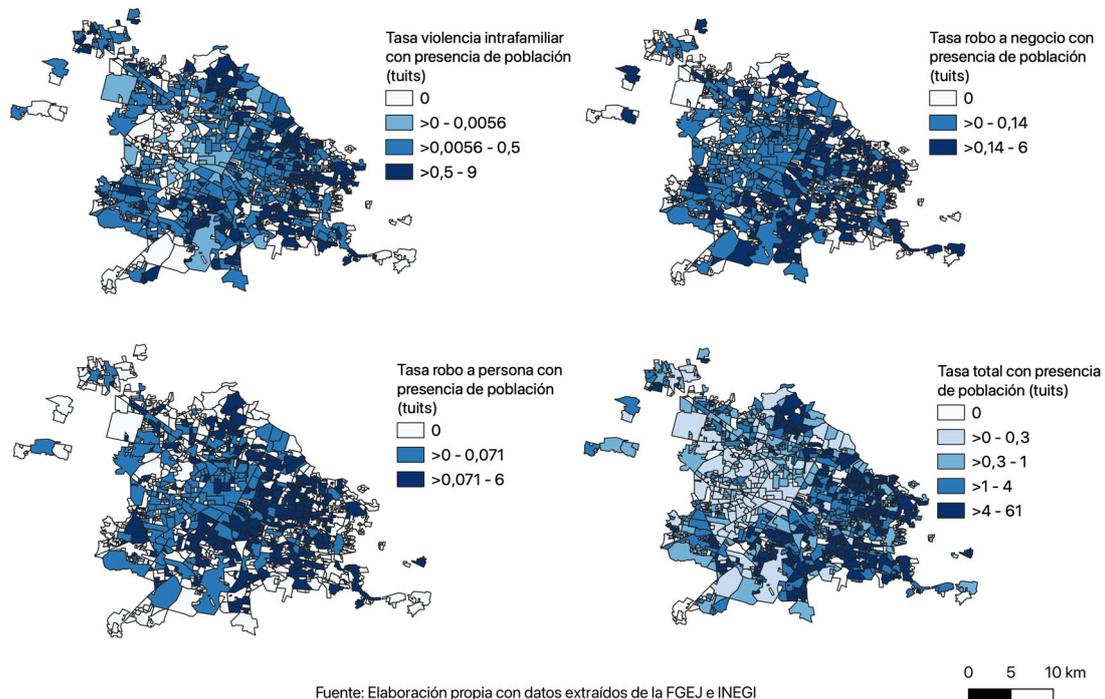


FIGURA 7.
TASA DE DELITO TOTAL Y DE CLASIFICACIONES EN ESTUDIO CALCULADAS CON PRESENCIA DE POBLACIÓN (TUIITS)



sencia de población representada como cantidad de tuits georreferenciados, se aprecia que las colonias con mayor tasa de delincuencia se presentan en el oriente y norte de la ZMG. En la figura 8 se puede apreciar la tasa de delincuencia utilizando los establecimientos y unidades económicas como denominador, en el cual se puede apreciar que también hay concentración de altas tasas en la región oriente y norte de la ciudad lo cual se asemeja a la figura 7, pero también ilustra colonias con altas tasas de delincuencia en el poniente de la ciudad.

Por último, en la figura 9, se puede apreciar la tasa de delincuencia calculada con la población residencial. Con este tipo de población las colonias con mayor tasa de delitos se concentran en el centro, poniente y sur de la ciudad, esto es claramente una diferencia con lo encontrado en las figuras 7 y 8, lo que representa una identificación distinta de la población en riesgo. Además, se puede apreciar que las concentraciones son menos claras y precisas, lo que dificulta la concentración de esfuerzos en temas de política de prevención y combate al delito.

Clusters Getis-Ord

A continuación, se muestran las áreas donde las distintas tasas de delito calculadas son significativas al

95%⁹ calculadas a través del estadístico Getis Ord para el total de delitos y tres tipos de delitos específicos: robo a persona, robo a negocio y violencia intrafamiliar. Las áreas de color rojo (alto) son *clusters* de colonias donde hubo altas tasas de delincuencia rodeadas de colonias con altas tasas de delincuencia, por otro lado, se puede apreciar los *clusters* bajos (en color azul), que representan colonias con bajas tasas de delincuencia rodeadas de colonias con bajas tasas de delincuencias.

De acuerdo con Malleson y Andresen (2016), los *clusters* de altas tasas de robo a persona basadas en presencia de población pueden interpretarse como colonias que son atractoras de crimen, ya que son áreas que cuentan con un mayor volumen de delincuencia del que se esperaría dado el tamaño de la población en riesgo, que en este caso es la presencia de población, es decir, son áreas donde los delincuentes no son oportunistas si no que van a cometer el crimen de manera premeditada. Asimismo, los *clusters* de altas tasa de crimen basadas en población residencial se consideran zonas de generadoras de crimen, ya que cuentan con mayor volumen de delincuencia del que se esperaría dado el tamaño de la población

9 Esto se trabaja con el estadístico Gi, explicado en el apartado de metodología

FIGURA 8.
TASA DE DELITO TOTAL Y DE CLASIFICACIONES EN ESTUDIO CALCULADAS CON PRESENCIA DE POBLACIÓN (ESTABLECIMIENTOS)

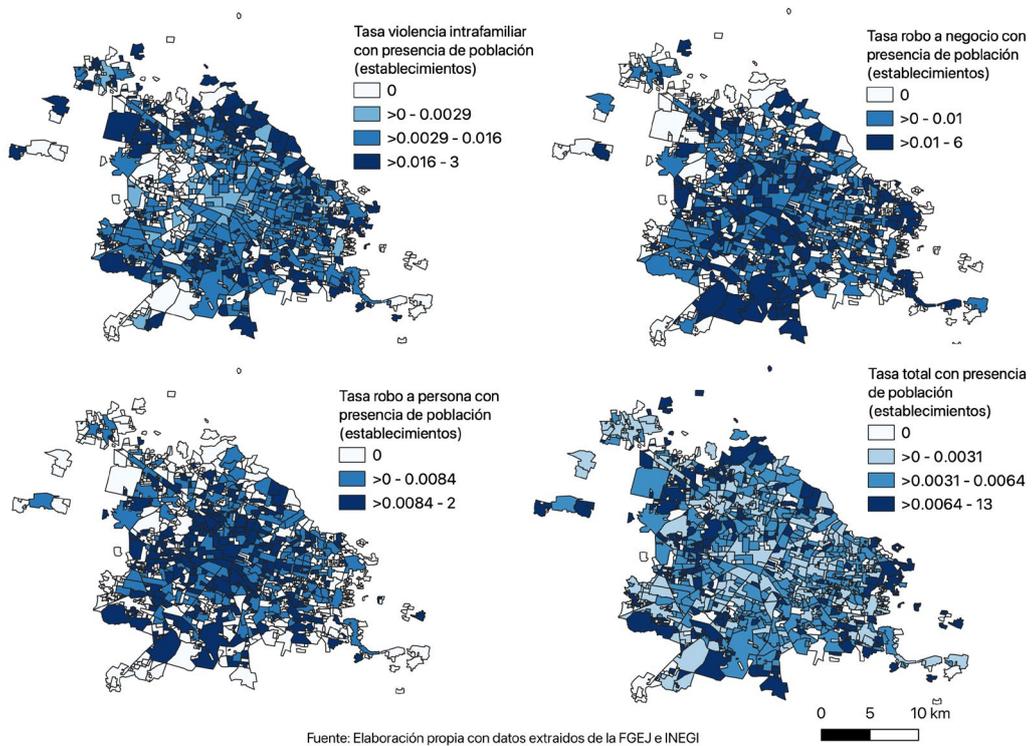
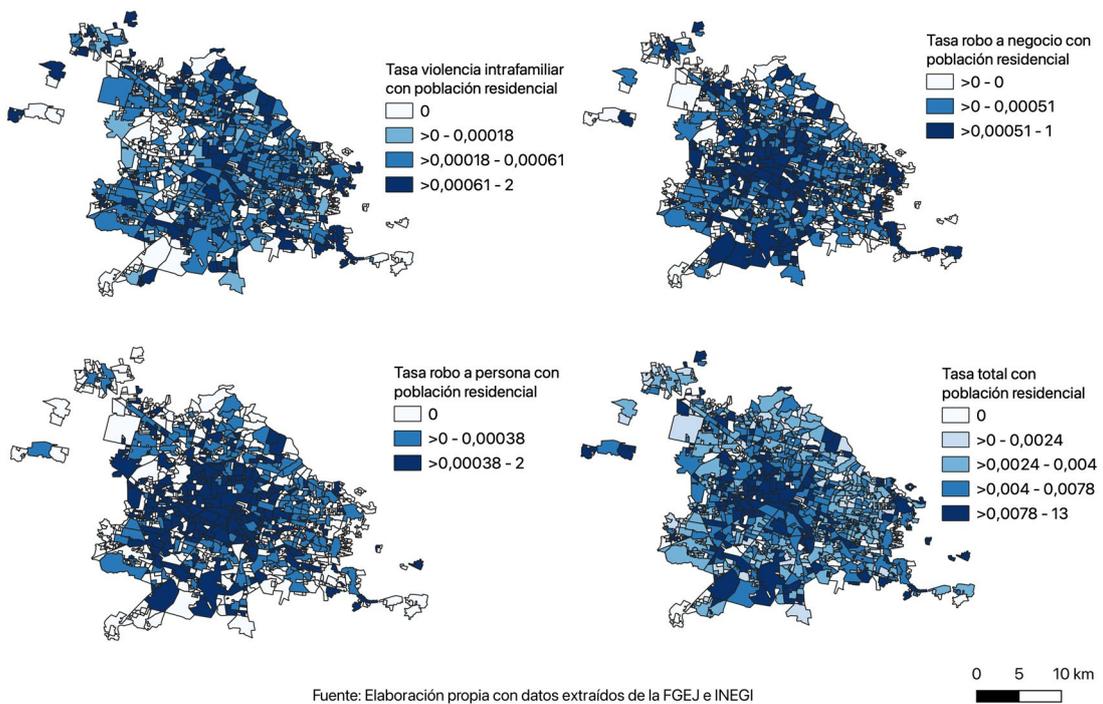


FIGURA 9.
TASA DE DELITO TOTAL Y DE CLASIFICACIONES EN ESTUDIO CALCULADAS CON POBLACIÓN RESIDENCIAL



en riesgo que es la población residencial, pudiéndose interpretar que en este caso los delincuentes no iban con la intención de cometer el crimen, sino que son delincuentes oportunistas.

Por lo anterior, únicamente en el caso del delito de robo a persona se presenta adicionalmente un análisis de zonas atractoras y generadores de delito, pues los otros delitos analizados no necesariamente involucran el movimiento de las víctimas del crimen.

a. Clusters Getis-Ord para tasas de robo a persona

En las figuras 10, 11 y 12 se observan los *clusters* (*high* y *low*) de las tasas de robo a persona. Se puede apreciar que existen áreas donde se ven tasas de robo a persona significativamente altas utilizando la presencia de población como denominador, pero no son significativas cuando se utiliza la población residencial como denominador. Si se hubiera utilizado solo la población residencial, estas áreas no se

FIGURA 10.
ESTADÍSTICOS GETIS ORD PARA TASAS DE DELITO CALCULADO CON PRESENCIA DE POBLACIÓN (TUITTS)

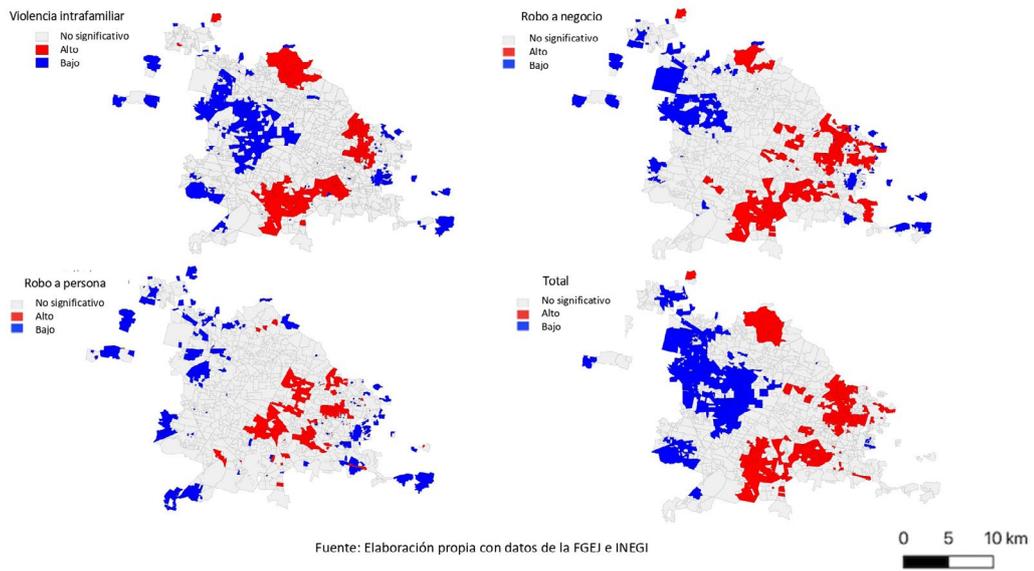


FIGURA 11.
ESTADÍSTICOS GETIS ORD PARA TASAS DE DELITO CALCULADO CON PRESENCIA DE POBLACIÓN (ESTABLECIMIENTOS)

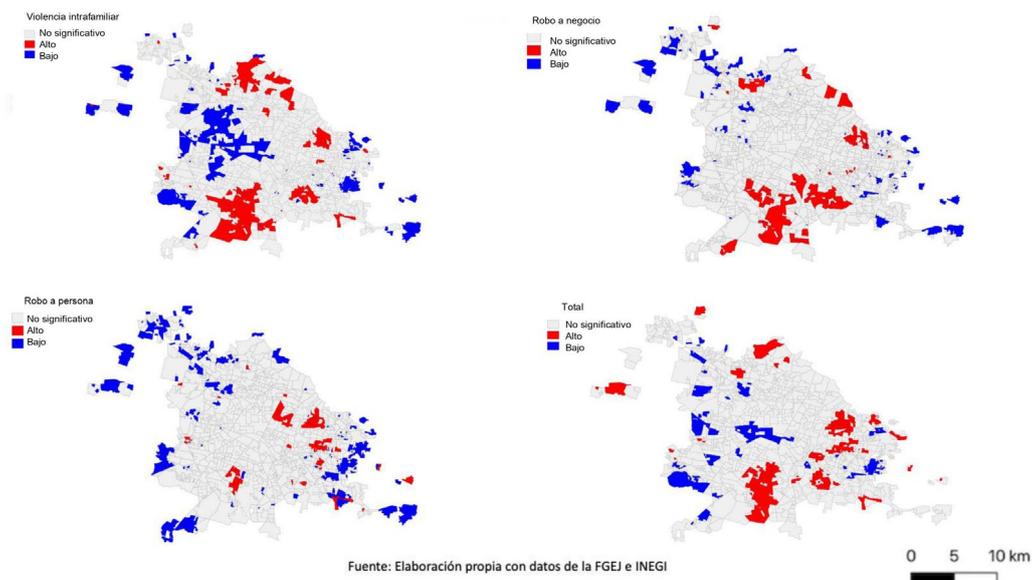
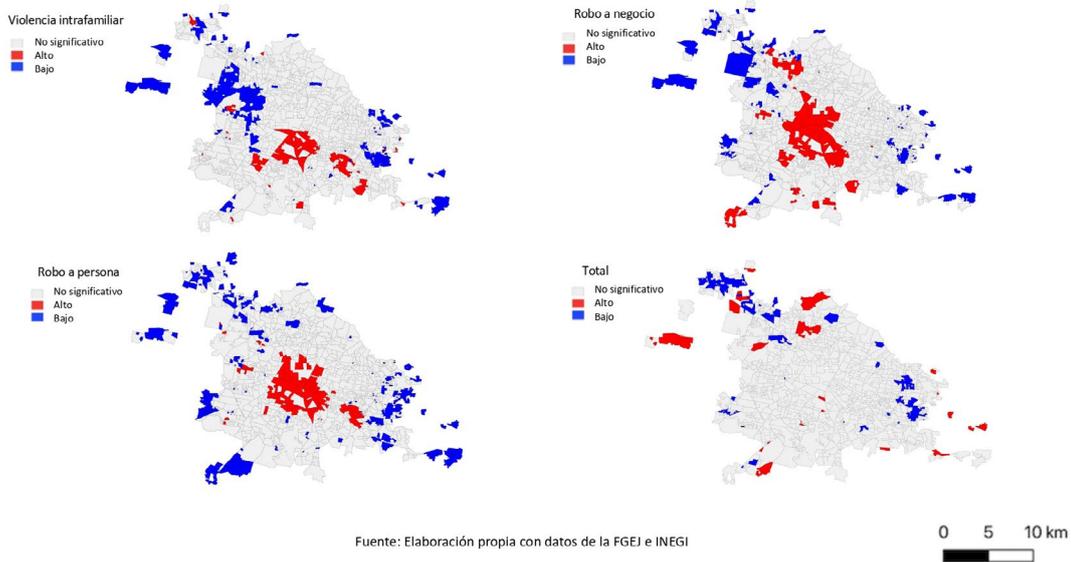


FIGURA 12.
ESTADÍSTICOS GETIS ORD PARA TASAS DE DELITO CALCULADO CON POBLACIÓN RESIDENCIAL



considerarían como *hotspots* de altas tasas de robo a persona (Malleison & Andresen, 2016). Es importante recordar que el robo a persona obtuvo mayor correlación con los tuits de entre los distintos tipos de población utilizados.

Como se puede observar en la figura 13, existen *clusters* que coinciden entre la tasa de robo a persona basada en presencia de población (tuits o establecimientos) y tasa de robo a persona basada en población residencial. Esto permite suponer que son áreas predominantemente residenciales que ofrecen comodidades para la movilidad de las personas. Estas colonias se pueden clasificar como áreas mixtas, es decir, son generadoras y atractoras al mismo tiempo, pues atraen a un gran número de personas

que no tienen fines delictivos, así como lugares que se usan específicamente por los criminales (Malleison & Andresen, 2016).

Para el robo a personas se aprecia que las áreas coincidentes entre *clusters* Getis basadas en tuits y *clusters* Getis basados en población residencial se encuentran en la zona centro. En cambio, las áreas coincidentes entre *clusters* Getis de tasas basadas en los establecimientos y los *clusters* Getis basados en población residencial son prácticamente imperceptibles.

La figura 14 muestra la cantidad de establecimientos por giro comercial por km² que intersectan con los *clusters* generados con las distintas tasas de robo a

FIGURA 13.
ÁREAS COINCIDENTES ENTRE CLUSTERS GETIS-ORD

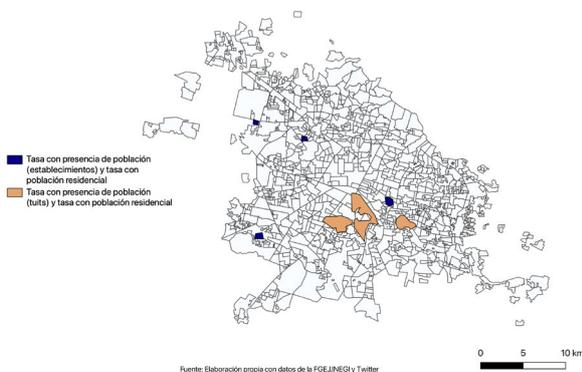
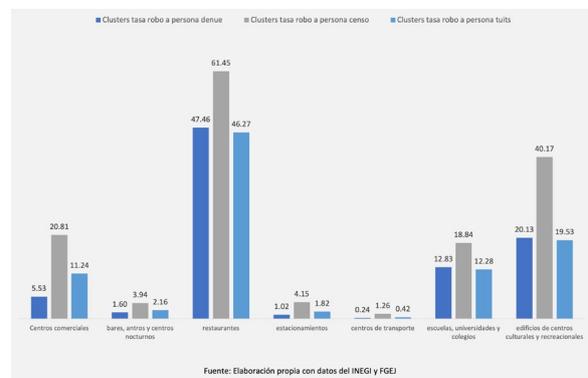


FIGURA 14.
CANTIDADES DE ESTABLECIMIENTOS POR KM2 QUE INTERSECTAN CON CLUSTERS DE TASA DE ROBO A PERSONA



persona, esto confirma que los *clusters* de presencia de población son colonias mayormente de atracción de crimen ya que hay pocos generadores (restaurantes, bares) por lo que podrían existir atractores que provocan los crímenes.

b. *Clusters* Getis-Ord para tasas de robo a negocio

Es importante recordar que la correlación entre el robo a negocio y los establecimientos y unidades económicas fue la más alta de entre dicha variable y los distintos tipos de crimen. En las figuras 10, 11 y 12 también se pueden apreciar los *clusters* generados por las tasas de robo a negocio con los distintos tipos de población. Los *clusters* de tasas de robo a negocio basadas en tuits y en el DENUÉ se encuentran en el oriente de la ZMG, mientras que los *clusters* de tasas de robo a negocio basados en población residencial se concentran en colonias del centro y noroeste de la ZMG. Es decir, también hay diferencias claras en el cálculo utilizando la población residencial y la presencia de población.

c. *Clusters* Getis-Ord para tasas de violencia intrafamiliar

De entre todos los tipos de delito, la violencia intrafamiliar fue el delito que mayor correlación obtuvo con la población residencial, lo cual tiene sentido teniendo en cuenta que gran parte de estas conductas ocurren al interior del hogar y no se relacionan con la movilidad de las personas.

Nuevamente, las figuras 10, 11 y 12 muestran que los *clusters* generados por las tasas basadas en la presencia de población se ubican en colonias del norte, sur y oriente de la ciudad, mientras que los *clusters* generados con las tasas basadas en población residencial se encuentran en el centro y sureste de la ZMG. En este caso, dada la naturaleza de este delito y la correlación con la población residencial, es de suponer que la concentración mejor reflejada sea la de la figura 12 en el que se utiliza la población residencial, pero como ocurre al igual que en los otros delitos, las diferencias utilizando las diferentes medidas son significativas.

DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos dan lugar a varios puntos de reflexión. En primer lugar, en el desarrollo de este trabajo se comprobó que los tuits otorgan información adicional que explica el comportamiento del delito en una unidad de análisis pequeño, como lo es la

colonia o barrio. Si bien existen importantes limitaciones de la utilización de esta red social que hacen que nuestros resultados deben tomarse con precaución, siendo la primera de ellas la posibilidad de infrarrepresentación que puede suponer un sesgo al relacionarse con el poder adquisitivo, la edad o el nivel de educación y no solo la presencia de población, puede ser una medida que con el paso del tiempo vaya adquiriendo mayor relevancia para realizar estimaciones de ciertos delitos de manera más precisa y periódica. Ante esto, sabemos que Twitter es mayormente usada por jóvenes; de acuerdo con la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (INEGI, 2020b) del INEGI, el 66% de los usuarios de Twitter en México están en el rango de edad de entre 13 y 32 años; sin embargo, hay que recordar que también son los jóvenes los que más son víctimas de delito y los que más denuncian (INEGI, 2020a), por ello, tal y como señalan Malleon y Andersen (2014) es posible que los tuits sean una buena herramienta para identificar específicas subpoblaciones en riesgo, como los jóvenes, pero que con el tiempo cada vez representen a mayor parte de la población.

Otro punto de discusión relacionado a los datos de Twitter es que solo el 9,7% de los usuarios de internet usan esta red social en México debido a la brecha digital existente (INEGI, 2020b). Sin embargo, la correlación encontrada entre el grado de marginación y los tuits fue negativa y significativa pero muy baja, lo cual disminuye el posible sesgo introducido por este hecho y, por otro lado, la evolución digital ha generado que cada vez más personas utilicen internet, es por esto que si bien existen inconvenientes actuales de la utilización de estos datos, esta herramienta será aún mejor y más extendida en el futuro ya que se espera que la brecha digital vaya disminuyendo paulatinamente.

Otra limitante es el alto nivel de concentración de los tuits en pocas colonias. En este estudio se recolectaron 66,087 tuits georreferenciados, de los cuales la mitad se concentran en 22 colonias de la ZMG. Sin embargo, a pesar de ello, y como se enfatiza más adelante, el análisis ha comprobado que los tuits pueden ser una alternativa para el cálculo de medidas alternativas del delito.

A pesar de las limitaciones presentadas en este trabajo, como se comentó en la introducción, existen enormes ventajas de representar y utilizar la presencia de población a través de tuits y establecimientos comerciales, las cuales se pueden resumir en que los tuits son una fuente de información gratuita, las encuestas intercensales no son significativas a nivel co-

lonia y, tanto los tuits como el DENUÉ tienen ventaja en la periodicidad, ya que los censos poblacionales se realizan cada diez años en México; pero además, desde un punto de vista espacial, el presente trabajo aporta evidencia de la conveniencia de utilizar otros indicadores para determinar la población en riesgo y que, además, es deseable utilizar diferentes indicadores de población en riesgo dependiendo del delito que se pretenda analizar. Un área de oportunidad en la cual se podrían reducir las limitaciones presentadas por los datos extraídos de Twitter en el tema de la georreferencia, es realizar trabajos donde se utilice esta variable como *proxy* a la presencia de población a nivel localidad y no a nivel colonia, sin embargo, para poder llevar a cabo estos trabajos se tendría que contar con datos de delitos a nivel localidad diarios, caso que no sucede para las localidades que integran la ZMG.

En los resultados se aprecia que se cumple la hipótesis planteada ya que para distintos tipos de delito se identificó distinto tipo población en riesgo: para el caso del robo a persona se encontró mayor correlación con la presencia de población basada en tuits, para el caso de robo a negocio se encontró mayor correlación con la presencia de población basada en establecimientos y para el caso de violencia intrafamiliar se encontró mayor correlación con la población residencial. Esto justifica que la presencia de población en el ambiente es una variable de utilidad para programas y políticas centrados en combatir el crimen en la ZMG, ya que es claro que la población residencial no siempre es la más adecuada para identificar la población en riesgo, y más aún para análisis desagregados de la criminalidad como lo es un análisis a nivel colonia. A diferencia de Malleson y Andresen (2016), donde se encontró una mayor correlación entre robo a persona y la variable de establecimientos, en este trabajo se encontró mayor correlación entre robo a persona y tuits. Sin embargo, al igual que en Malleson y Andresen (2016) y Malleson y Andresen (2015) se encontró que la correlación entre los delitos de robo y la población residencial fue la más baja de entre los distintos tipos de población, excepto en el caso de violencia intrafamiliar, lo cual era algo planteado en la hipótesis ya que es un tipo de delito que acontece en el núcleo familiar.

Creando *clusters* con el estadístico Gi y basados en Malleson y Andresen (2016) se pudieron identificar zonas atractoras de crimen y generadoras de crimen para el caso de robo a persona. Las zonas atractoras

de crimen se identificaron como los *clusters* High¹⁰ de tasas de delito basadas en la presencia de población, ya que son zonas donde los delitos son más altos si se compara con la población que transita por dichas zonas, esto significa que los crímenes que se generan no son oportunistas, sino que el delincuente va premeditado a cometer el crimen. Por otro lado, las zonas generadoras de crimen se identificaron como los *clusters* High de tasas de delito basadas en población residencial, ya que los delitos son más altos si se compara con la población que vive en dichas zonas. Por último, las zonas coincidentes son consideradas zonas mixtas, es decir, son zonas tanto atractoras como generadoras de crimen. Estos hallazgos pueden constituir una herramienta poderosa en términos de prevención del delito ya que, por ejemplo, las estrategias de prevención situacional van exclusivamente dirigidas a disminuir las oportunidades del delito y los resultados del presente trabajo permiten ir más allá de esta presunción e identificar zonas en las que la oportunidad no es el factor principal en la comisión del delito de robo a persona.

Este trabajo propone alternativas de medición que permiten informar de una manera más precisa a los tomadores de decisiones en materia de seguridad ciudadana el lugar donde se concentran los delitos, así como las colonias que son atractoras y generadoras de crimen. Con ello, se establecieron tres medidas de la población en riesgo que permitieron establecer patrones espaciales distintos y por lo tanto una mayor precisión en la estimación de las tasas delictivas. Es por esto que este trabajo tiene implicaciones de política pública, ya que puede servir a los tomadores de decisiones y cuerpos policíacos a enfocar de manera más eficiente los esfuerzos de combate al crimen en la ZMG, no solo al identificar zonas de riesgo con *clusters* de altas tasas de delito basadas en población residencial, sino también zonas de riesgo con *clusters* de altas tasas de delito basadas en la presencia de población en el ambiente. Para ello, se han utilizado algunas medidas que, si bien no están exentas de limitaciones, sí pueden representar un enorme potencial para el futuro de la política criminal y la prevención del delito.

REFERENCIAS

Asociación Mexicana del Internet (2016). *Estudio de Hábitos de los Usuarios de Internet en México*. Ciudad de México: AMIPCI.

10 Colonias con altas tasas de delito rodeadas de colonias con altas tasas de delito

- Andresen, M. A. (2006). Crime measures and the spatial analysis of criminal activity. *British Journal of Criminology*, 46(2), 258-285. <https://doi.org/10.1093/bjc/azi054>
- Andresen, M. A. (2011). The ambient population and crime analysis. *The Professional Geographer*, 63(2), 193-212. <https://doi.org/10.1080/00330124.2010.547151>
- Andresen, M. A. & Jenion, G. W. (2008). Crime prevention and the science of where people are. *Criminal Justice Policy Review*, 19(2), 164-180. <https://doi.org/10.1177/0887403407311591>
- Andresen, M. A. & Linning, S. J. (2012). The (in)appropriateness of aggregating across crime types. *Applied Geography*, 35(1-2), 275-282. <https://doi.org/10.1016/j.apgeog.2012.07.007>
- Blank, G. (2017). The Digital Divide Among Twitter Users and Its Implications for Social Research. *Social Science Computer Review*, 35(6), 679-697. <https://doi.org/10.1177/0894439316671698>
- Brantingham, P. & Brantingham, P. (Eds.). (1981). *Environmental Criminology*. Sage Publications.
- Brantingham, P. & Brantingham, P. (1995). Criminality of place. *European Journal on Criminal Policy and Research*, 3(3), 5-26. <https://doi.org/10.1007/BF02242925>
- Brantingham, P. & Brantingham, P. (2008). Crime pattern theory. En R. Wortley & L. Mazerolle (Eds.) *Environmental Criminology and Crime Analysis* (pp. 78-93). Willan Publishing. <https://doi.org/10.4324/9780203118214>
- Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación & Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación. (2020). *Medición de Movilidad usando Facebook, Google y Twitter*. Consejo Nacional de Ciencia y Tecnología.
- Chan, J., & Bennett M. L. (2015). Is Big Data challenging criminology? *Theoretical Criminology*, 20(1), 21-39. <https://doi.org/10.1177/1362480615586614>
- Cheng, T. & Chen, T. (2021). Urban Crime and Security. En W. Shi, M. F. Goodchild, M. Batty, K. Mei-Po & Z. Anshu Zhang (Eds.), *Urban Informatics* (pp. 213-228). Springer.
- Fiscalía General del Estado de Jalisco (2021). Número de delitos registrados en carpetas de averiguación de la Fiscalía de Jalisco por año de 2016 a agosto 2021. <https://datos.jalisco.gob.mx/dataset/numero-de-delitos-registrados-en-carpetas-de-averiguacion-de-la-fiscalia-de-jalisco-de-2016>
- Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 588-608. <https://doi.org/10.2307/2094589>
- Getis, A., & Ord, J. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3), 189-206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>
- Hawelka, B., Sitko, I., Beinát, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260-271. <https://doi.org/10.1080/15230406.2014.890072>
- Instituto Nacional de Estadística y Geografía. (2010). *Censo de Población y Vivienda*. Aguascalientes: Instituto Nacional de Estadística y Geografía.
- Instituto Nacional de Estadística y Geografía (2017). *Directorio Estadístico Anual de Unidades Económicas (DENUE)*. Aguascalientes: Instituto Nacional de Estadística y Geografía.
- Instituto Nacional de Estadística y Geografía. (2020a). *Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) 2020*. Aguascalientes: Instituto Nacional de Estadística y Geografía.
- Instituto Nacional de Estadística y Geografía. (2020b). *Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2020*. Aguascalientes: Instituto Nacional de Estadística y Geografía.
- Kinney, J. B., Brantingham, P. L., Wuschke, K., Kirk, M. G., & Brantingham, P. J. (2008). Crime attractors, generators and detractors: Land use and urban crime opportunities. *Built Environment*, 34(1), 62-74. <https://doi.org/10.2148/benv.34.1.62>
- Lubbe, S., Filzmoser, P., & Templ, M. (2021). Comparison of zero replacement strategies for compositional data with large numbers of zeros. *Chemometrics and Intelligent Laboratory Systems*, 210, 104248. <https://doi.org/10.1016/j.chemolab.2021.104248>
- Malleson, N., & Andresen, M. A. (2014). The impact of using social media data in crime rate calculations: shifting hot spots and changing spatial pat-

- terns. *Cartography and Geographic Information Science*, 42(2), 112-121. <https://doi.org/10.1080/15230406.2014.905756>
- Malleson, N., & Andresen, M. A. (2015). Spatio-temporal crime hotspots and the ambient population. *Crime Science*, 4(1), 1-8. <https://doi.org/10.1186/s40163-015-0023-8>
- Malleson, N., & Andresen, M. A. (2016). Exploring the impact of ambient population measures on London crime hotspots. *Journal of Criminal Justice*, 46, 52-63. <https://doi.org/10.1016/j.jcrimjus.2016.03.002>
- Maltz, M. D. (1999). *Bridging gaps in police crime data*. DIANE Publishing.
- Osorio, J. (2020). Análisis de los patrones espacio-temporales de eventos a partir de datos de Twitter: el caso de la World Pride 2017 en Madrid. *Estudios Geográficos*, 81(288), e032. <https://doi.org/10.3989/estgeogr.202047.027>
- Pramanik, M. I., Lau, R. Y. K., Yue, W. T., Ye, Y., & Li, C. (2017). Big data analytics for security and criminal investigations. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(4), e1208. <https://doi.org/10.1002/widm.1208>
- Ristea, A., Langford, C., & Leitner, M. (2017, August). Relationships between crime and Twitter activity around stadiums. *2017 25th International Conference on Geoinformatics* (pp. 1-5). IEEE. Doi: 10.1109/GEOINFORMATICS.2017.8090933
- Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública. (2017). Datos abiertos de incidencia delictiva. Ciudad de México: Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública.
- Snaphaan, T., & Hardyns, W. (2019). Environmental criminology in the big data era. *European Journal of Criminology*, 18(5), 147737081987775. <https://doi.org/10.1177/1477370819877753>
- Traunmueller, M., Quattrone, G., & Capra, L. (2014, November). Mining mobile phone data to investigate urban crime theories at scale. En L. Aiello & D. McFarland (Eds.) *Social Informatics* (pp. 396-411). Springer. https://doi.org/10.1007/978-3-319-13734-6_29
- Trinidad, A., San Juan, C., & Vozmediano, L. (2019). Escenarios de delincuencia juvenil en el ámbito urbano: una perspectiva situacional. *Revista Criminología*, 61(2), 9-24.
- Wang, X., Gerber, M. S., & Brown, D. E. (2012). Automatic crime prediction using events extracted from twitter posts. En S. Yang; A. Greenberg & M. Ensley (Eds.) *Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 231-238). Springer. https://doi.org/10.1007/978-3-642-29047-3_28
- Wang, H., Kifer, D., Graif, C., & Li, Z. (2016, August). Crime rate inference with big data. En *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 635-644). <https://doi.org/10.1145/2939672.2939736>
- Williams, M. L., Burnap, P., & Sloan, L. (2017). Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns. *The British Journal of Criminology*, 57(2), 320-340. <https://doi.org/10.1093/bjc/azw031>
- Wortley, R., & Mazerolle, L. (2008). Environmental criminology and crime analysis: situating the theory, analytic approach and application. En R. Wortley & L. Mazerolle (Eds.) *Environmental criminology and crime analysis* (pp. 1-18). Willan Publishing. <https://doi.org/10.4324/9780203118214>